

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-210029

(43) 公開日 平成10年(1998) 8月7日

(51) IntCl.⁶

H 0 4 L 12/18

識別記号

F I

H 0 4 L 11/18

審査請求 未請求 請求項の数12 OL (全 18 頁)

(21) 出願番号 特願平10-6494

(22) 出願日 平成10年(1998) 1月16日

(31) 優先権主張番号 08/785625

(32) 優先日 1997年1月17日

(33) 優先権主張国 米国 (US)

(71) 出願人 596077259

ルーセント テクノロジーズ インコーポ
レイテッド

Lucent Technologies
Inc.

アメリカ合衆国 07974 ニュージャージ
ー、マレーヒル、マウンテン アベニュー
600-700

(72) 発明者 シャオカン チェン

アメリカ合衆国、07751 ニュージャージ
ー、モーガンビル、パークレー コート
32

(74) 代理人 弁理士 三俣 弘文

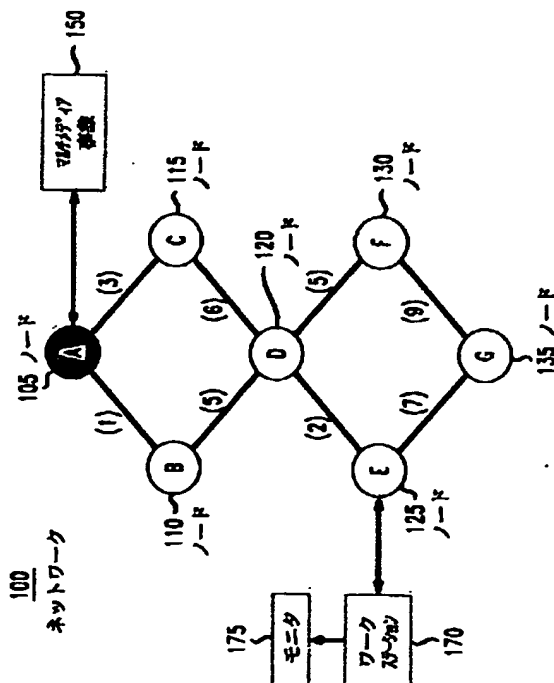
最終頁に続く

(54) 【発明の名称】 マルチキャスト情報を分配するソースノードから開始されるマルチキャスト接続に参加する方法

(57) 【要約】

【課題】 安価かつ高効率のマルチキャストグループを
探しだし、これに参加しかつこれから離れるためのマル
チキャストプロトコルを提供する。

【解決手段】 例えばATM交換機のような複数の通信
ノード(105, 110, ...)から形成される通信ネ
ットワーク(100)におけるマルチキャスト情報の分
配は、マルチキャストの開始ノードへのマルチキャスト
接続に参加するための要求をルーティングするための効
率のよいメカニズムおよび要求ノードをマルチキャスト
接続に接続するための効率のよいメカニズムを提供する
ことにより強化される。



Best Available Copy

【特許請求の範囲】

【請求項1】 マルチキャスト情報を分配するソースノードから開始されるマルチキャスト接続に参加する方法であって、前記マルチキャスト接続が、複数の通信ノード(105, 110, ...)からなるネットワーク(100)中で確立されるものにおいて、

前記ノードのうちの一つにおいて、前記マルチキャスト情報を受信するための要求の受信に応じて、選択されたバスからツリー中で識別された多数の他のノードの内のそれぞれに対して形成される少なくとも一つのルーティングツリーを構成し、かつ前記情報を識別するメッセージを前記ルーティングツリーにしたがって前記選択されたノードへ送信するステップと、

前記一つのノードにおいて、前記ソースノードを識別された情報の開始ノードとして識別するメッセージの受信に応じて、前記識別された情報のマルチキャスト分配に参加するための要求を含むメッセージを送信するステップとを有することを特徴とする方法。

【請求項2】 前記ソースノードにおいて、前記識別された情報のマルチキャスト分配に参加するための要求を含むメッセージの受信に応じて、肯定応答メッセージを前記メッセージの開始ノードへ返信するステップをさらに含むことを特徴とする請求項1記載の方法。

【請求項3】 前記ソースノードにおいて、前記識別された情報のマルチキャスト分配に参加するための要求を含むメッセージの受信に応じて、前記ソースノードに根を有するネットワークバスツリーを構成し、前記メッセージを発するノードへ拡張するステップと、前記マルチキャスト情報を前記ソースノードに根を有するバスツリーを経由して供給するステップとをさらに含むことを特徴とする請求項1記載の方法。

【請求項4】 前記マルチキャストツリーが、前記ソースノードと前記メッセージを発するノードとの間の中間ノードを含み、

前記中間ノードにおいて、前記マルチキャスト情報に対する要求を含むメッセージを前記ノードのうちの別の一つから受信することに応じて、前記ノードのうちの前記別の一つへ、前記中間ノードが前記マルチキャスト情報を供給することができることを表示する応答メッセージを返信するステップと、

前記ノードの内の前記別の一つにおいて、マルチキャスト情報の要求に対する全ての応答メッセージを集めて、応答メッセージを返信するどのノードが、前記ノードのうちの前記別の一つに最も近いかを決定するステップと、

前記マルチキャストに参加するためのメッセージを、応答メッセージを前記ノードのうちの前記別の一つへ返信したノードのうちの決定された最も近い一つへ送信するステップとをさらに含むことを特徴とする請求項3記載の方法。

【請求項5】 前記ノードのうちの最も近い一つにおいて、参加メッセージの受信に応じて、かつその後のソースノードに根を有するバスツリーを経由するマルチキャスト情報の受信に応じて、受信したマルチキャスト情報のコピーを前記ノードの内の前記別の一つへ送信するステップをさらに含むことを特徴とする請求項4記載の方法。

【請求項6】 前記中間ノードが、前記ノードの内の前記最も近い一つであり、

10 前記中間ノードにおいて、前記一つのノードから前記ソースノードに向けられたメッセージを受信し、前記一つのノードへのマルチキャスト情報の送信終了を要求することに応じて、前記一つのノードのマルチキャスト情報の送信を終了し、前記ノードのうちの前記別の一つに対して受信されるように、前記マルチキャスト情報を供給することを継続するステップをさらに含むことを特徴とする請求項5記載の方法。

【請求項7】 前記中間ノードが、前記ノードのうちの前記最も近い一つであり、

20 前記中間ノードにおいて、前記ノードのうちの前記別の一つから前記ソースノードへ向けられたメッセージを受信し、前記ノードのうちの前記別の一つへのマルチキャスト情報の送信終了を要求することに応じて、前記ノードのうちの前記別の一つへのマルチキャスト情報の送信を終了させ、前記マルチキャスト情報に対する他の受信ノードが無いことに応じて、前記受信された終了メッセージを前記ソースノードへ送信するステップをさらに含むことを特徴とする請求項5記載の方法。

【請求項8】 マルチキャスト情報を分配するソースノードから開始されたマルチキャスト接続に参加するための方法であって、前記マルチキャスト接続が、複数の通信ノードからなるネットワークにおいて確立されるものにおいて、

一つのネットワークノードにおいて、前記ネットワークノードのうちの第1および第2の別の一つから、マルチキャストに参加するためのそれぞれのメッセージ要求を受信することに応じて、所定のパラメータの関数として、第1および第2のノードのうちのどれが、最初にマルチキャスト接続へ接続されるべきかを決定し、待機メッセージを前記第1および第2のノードのうちの前記他の一つへ送信するステップと、

40 前記第1および第2のノードのうちの前記一つから受信したメッセージを、そのメッセージ中に含まれるバスを経由して送るステップと、

マルチキャスト情報を前記ソースノードに根を有するバスを経由してソースから受信することに応じて、前記マルチキャスト情報を前記第1および第2のノードのうちの前記一つへ供給するステップとを有することを特徴とする方法。

50 【請求項9】 前記第1および第2のノードのうちの前

記別の一つにおいて、前記待機メッセージの受信に応じて、所定時間待機するステップと、

前記時間の満了時点において、マルチキャスト情報の受信を要求するメッセージを再度送信するステップとをさらに含むことを特徴とする請求項8記載の方法。

【請求項10】 前記ネットワークノードにおいて、前記第1および第2のノードのうちの前記別の一つからの要求メッセージの受信に応じて、前記マルチキャスト情報を、それが前記ソースから受信した時に、前記第1および第2のノードのうちの前記別の一つへ供給するステップをさらに含むことを特徴とする請求項9記載の方法。

【請求項11】 前記第1および第2のノードのうちの前記一つのノードから、前記ソースノードに向けられたメッセージを受信し、前記第1および第2のノードのうちの前記一つへ前記マルチキャスト情報の送信終了を要求することに応じて、マルチキャスト情報を前記ノードへ送信し、前記マルチキャスト情報をそれが前記第1および第2のノードのうちの前記別の一つへ受信されたときに、供給し続けるステップをさらに含むことを特徴とする請求項9記載の方法。

【請求項12】 前記ネットワークノードにおいて、前記第1および第2のノードのうちの前記一つのノードから、前記ソースノードに向けられたメッセージを受信し、前記第1および第2のノードのうちの前記一つへ前記マルチキャスト情報の送信終了を要求することに応じて、前記ノードへのマルチキャスト情報の送信を終了させ、前記マルチキャスト情報に対する他の受信ノードが無いことに応じて、終了を要求する前記受信されたメッセージを、前記ソースノードへ送信するステップをさらに含むことを特徴とする請求項9記載の方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、データネットワークに係り、特に、特定の情報を受信するマルチキャストグループを探し出し、これに参加し、かつこれを離れるためのマルチキャストプロトコルに関する。

【0002】

【従来の技術】現在、例えば通常のワークステーションまたはPCのようなマルチメディア端末と関連づけられたユーザは、デジタルネットワークにおいて、パケット交換機のようなノードにより提供されているレクチャ、オーディオ/ビデオ会議、放送される音声のような特定の事象に参加する要求を、その端末により入力することができる。

【0003】この事象は、典型的には、あるタイプの識別子、即ちグループ識別子（例えば、通常の電話ネットワークにおける800番号（受信者負担電話番号）または会議ブリッジング番号）と関連づけられているので、事象に参加することを望むユーザは、そのユーザが彼／

彼女の端末に与える要求において事象を特定することができる。そして、ユーザの端末は、この要求を、デジタルネットワーク内の関連するサービス中のノードに送る。

【0004】グループ識別子は、事象と関連づけられた特定の情報に興味のあるユーザの集合を表すことができる。グループ識別子は、特定のユーザを識別するために使用される識別子とは異なる。典型的に、グループ識別子に関連づけられたメンバーシップは、ダイナミックで有り得るので、メンバーは、いかなる時においてもそのグループに入りかつグループから離れることができる。

【0005】また、同じ物理的な位置にあってもよく、そうでなくてもよいマルチキャストグループメンバーの物理的な位置についての制限はない。ネットワークは、そのグループに関する情報を維持し、典型的には、アクティブなメンバーの接続を追跡するために使用されるネットワークディレクトリを構成することにより、これを行う。即ち、サービス中のノードは、グループ識別子をネットワークディレクトリに与える。

【0006】

【発明が解決しようとする課題】このディレクトリは、参加することができる最も近いノードのアドレスを返信する。そして、サービス中のノードは、マルチキャストグループに参加する要求を識別されたノードに送信する。マルチキャストグループに参加し、またはこれを離れるノードは、メンバーシップ情報を更新することができるように、ディレクトリを通知しなければならない。あいにく、メンバーシップはあまりにも頻繁に変化し得るので、そのようなディレクトリの管理は、費用がかかりかつ非効率になり得る。

【0007】また、ネットワークは、ネットワーク中にマルチキャストサーバを構成することによりそのようなメンバーシップを維持し得る。そして、ユーザにより送信されるデータは、まずサーバに送られて、サーバは、この情報をマルチキャストグループ中の他の全てのメンバーに送る。このアプローチは、具現化が単純であるように見えるが、サーバに、事象中の単一点故障を生じさせる可能性がある。また、このアプローチは、ユーザデータがサーバに送信されなければならないので、ネットワーク資源を有効に使用しない。

【0008】

【課題を解決するための手段】本発明は、ネットワーク中のいずれのノードも、マルチキャストグループを自動的に探し出し、これに参加しかつこれを離れることが可能である分散形式で、マルチキャストグループ情報を維持するために、データネットワークのための効率のよいかつ標準化可能なメカニズムを提供する。このメカニズムは、単一のノードがマルチキャストグループの他の参加者に関する完全な情報を維持しなければならないように分散される。

【0009】特に、本発明の例示的な実施形態によれば、ツリー中で識別された多数の他のノードのそれぞれへの選択されたパスから形成された少なくとも1つのルーティングツリーを構成し、かつマルチキャストグループを識別する発見メッセージをルーティングツリーに従って選択されたノードへ送信することにより、マルチキャストグループに参加するための要求を入力する。マルチキャストグループに接続された最も近いノードを識別するメッセージの受信に応じて、要求を発しているノードは、マルチキャストグループに参加するために、参加メ

ッセージを識別されたノードへ単純に送信する。

【0010】

【発明の実施の形態】我々の新規なプロトコルは、我々が「参加者により開始される参加 (Participant-Initiated Join)」と呼ぶもの、即ちノードがマルチキャストグループに参加するための要求を発し、そうすることにおいて、マルチキャストに参加するために使用されるべきネットワークパスを決定するものをサポートする。第1の段階は、既にマルチキャストツリー状にあるノードのアイデンティティを決定する要求を発するノードであ

るので、これを「発見 (Find)」段階と呼ぶ。
【0011】第2の段階は、要求を発するノードが、マルチキャストツリー上にすでにある最も近いノードに参加しようと試みるので、これを「参加 (Join)」段階と呼ぶ。発見段階において、マルチキャストグループに参加しようとするノードは、要求メッセージを生成し、要求を発しているノードに根を有する最も短いパスツリー上のその隣のノードの全てに要求メッセージを送信する。

【0012】このメッセージを受信する隣のノードは、このメッセージをその隣のノードへ、開始ノードに根を有する最も短いパスツリーのリンクに沿って、下流側へ送る。このプロセスは、要求メッセージが、既にマルチキャストツリー上にあるノードまたは開始ノードに根を有する最も短いパスツリーのリーフノード (即ち、求められているマルチキャスト情報を有するノード) のいずれかに到達するまで続けられる。

【0013】既にマルチキャストツリー上にある各ノードは、0にセットされたコストパラメータで、要求に回答する。マルチキャストツリー上にないリーフノードは、-1にセットされたコストパラメータで、要求に回答する。メッセージに対する全ての回答の受信において、ノードは、最も近いノードを決定し、最も近いノードから要求の開始ノードに向かって上流側に回答メッセージを送る。これを実行する前に、ノードは、このノードから最も近いノードへのコストを反映するために、コスト関数を更新する。

【0014】開始ノードは、全ての回答を受信し、最も近いノードを決定する。これで発見段階を終了する。参加段階において、要求を発しているノードは、最も近い

ノードへのパスを決定し、参加要求メッセージを送信する。この決定されたパスは、中間ノードが同じ決定をする必要がないように、メッセージ中に挿入される。求められる情報を有する最も近いノードは、参加肯定応答メッセージで応答し、パス中の全てのノードはマルチキャストツリーのブランチになる。

【0015】上述した例は、図1に示されており、求められるマルチキャスト情報は、マルチメディア事象150、例えば放送されるレクチャであると仮定される。マルチメディア事象を特徴づけるマルチメディア信号は、ノード105に供給され、ノード105は、データノード105、110、115、120、125、130および135から形成されるネットワーク100を経由して、そのような信号のコピーを受信するための要求を入力したユーザに対して、ビデオ信号をデジタル信号の形で供給するために、通常のビデオサーバとして動作し得る。

【0016】本発明の一実施形態において、例えば、各ノードは、通常のATM交換機である。ワークステーション170に関連づけられたユーザは、要求がマルチメディア事象に関連づけられた識別子を含むモニタ175上の事象を見るために、図示されていないキーボードにより通常の方法で要求を入力する。しかし、この要求は、マルチキャスト情報のソースのアイデンティティ (例えばアドレス) を含まない。

【0017】ワークステーション170は、要求の受信に応じて、Eノードとしても示される関連づけられたネットワークノード125により期待される形式に、要求を変換する。本発明の一側面として、ネットワーク100は、それぞれネットワーク100のノードにより供給される情報を識別するディレクトリを含まない。したがって、マルチキャスト情報を有するノードを探し出すために、ノード125は、他のネットワーク100のノードのそれぞれをプールすることになる。

【0018】しかし、それを実行する前に、ノード125は、選択されたパスから他のノードのそれぞれに対して形成されるツリーを構成し、選択は、例えばコスト、ホップの数等のような所定のパラメータに基づいてなされる。本発明の一実施形態において、コストは重み付け値により特徴づけられる。そのような重み付け値は、図1中のかっこ内に示されており、上述したように、少なくともコストパス/ルートを識別するために使用される。

【0019】例えば、ノード120 (D) がメッセージをノード105 (A) に送信する必要がある場合、ノード120はメッセージをおそらくノード110 (B) を経由して送信することになる。これは、そのパス中のリンクに関連づけられた重み付けの和が、ノード115 (C) を経由するパス中のリンクに関連づけられた重み付けの和よりも小さい、即ち、 $(5) + (1) < (6)$

+(3)であるからである。

【0020】したがって、既知の技術に従って、ノード125は、発見段階において、要求されるマルチキャスト情報、即ち放送される事象150を有するノードを探し出すために、バスツリーを構成し、メッセージのルーティングを制御するために、そのツリーを使用する。そのようなツリーの一例が図2に示されている。したがって、ノード120は、発見メッセージをツリーに従ってノード135(G)およびノード120(D)へ送信する。

【0021】この発見メッセージは、とりわけメッセージ開始ノードのアドレス、求められる情報に関連づけられた識別子、求められる情報を有するノードのアイデンティティに対する要求を含む。このメッセージは、構成されるバスツリーを特徴づける情報も含み得る。したがって、メッセージの受信に応じて、ノード135は、-1にセットされたコストパラメータとともに、応答メッセージを返信する。

【0022】ノード120は求められる情報を有していないが、図2に示されているように、バスツリーは、メッセージがノード110、115および130へ個々にルーティングされなければならないことを示すので、ノード120はメッセージを放棄しない。同様に、メッセージがノード110に到達した場合、メッセージが求められる情報を有していない場合にも、ノードはそのメッセージを放棄せず、それをノード105(A)へ送る。したがって、ノード105は求められる情報を有していると仮定され、ノード105は、コストパラメータを0にセットして、ノード105のアイデンティティ(アドレス)を含む応答メッセージをノード125に返信することにより、受信されたメッセージに応答する。

【0023】応答メッセージの受信により、ノード125は、上述したように、参加段階に入る。この段階において、ノード125は、ノード105への最短パスを構成し、識別された事象のマルチキャストに参加するための要求とともに、ノード125のアドレスを含む参加メッセージを送信する。

【0024】図2において、そのようなパスは、ノード115および120を経由するのではなく、ノード110および120を経由することになることがわかる。ノード105は、参加肯定応答を返信し、放送される事象を決定されたパスを経由してノード125へ供給することを開始する。マルチキャストのパス中にあるノード110および120がマルチキャスト情報の受信を開始した時、ノード110および120は、そのような情報を有することをそれらの各内部メモリに記録する。

【0025】この時点において、ノード130に関連づけられたユーザが、マルチキャストを受信し、その情報に対する要求を入力することも望むと仮定する。同様に、ノード130は、最小コストバスツリーを構成し、

図示しないツリーに従って、それ自身の発見メッセージを発する。このメッセージがノード120に到達した場合、情報を有していることを示す応答メッセージでノード130に応答する。

【0026】上述したように、ノード130は、応答メッセージを集めて、これらのメッセージから、求められる情報を有するノードのうちどれがノード130に最も近いかを決定する。この例において、所望の目的のために最も近いノードは、ノード120になる。したがって、ノード130は、参加メッセージをノード105へではなくノード120へ送信することにより、マルチキャストに参加することができる。そして、ノード120がそのような情報のパケットをノード110から受信した場合、ノード120は、コピーをノード125およびノード130へ送信する。

【0027】この時点において、ユーザがワークステーション170において、マルチキャストの受信を終了させるための要求を入力した場合、ノード125は分離メッセージを形成し、このメッセージをノード125に根を有する構成されたバスツリーに従ってノード120へ送信する。ノード120がこのメッセージを受信した場合、ノード120は、このメッセージをローカルメモリに格納し、ノード125へのマルチキャスト情報の供給を終了するが、ノード130により受信されるようにこの情報の供給を継続することになる。

【0028】放送される事象の終了に先だって、ノード130が分離メッセージを発する場合、このメッセージに応じて、かつマルチキャストに対する他の受信ノードが無いことに応じて、ノード120は、マルチキャストのノード120への供給を終了するために、メッセージをノード110に送信する。同様に、ノード110は、マルチキャスト情報に対する他の受信ノードがない場合、同様のメッセージをノード105へ送信する。

【0029】1つのノードが、ほぼ同じ時点で(同時に)、複数のノード、例えば2つのノードから発見メッセージを受信する状況が起こり得る。この状況を取扱い、かつ分散形プロトコルの正確さを保つために、1つのノード、例えばノード125は、発見メッセージに時間スタンプを付加する。

【0030】この方法において、1つのノード、例えばノードDが、ノード125および130から同時に発見メッセージを受信した場合、ノードDは、早い時間スタンプを有する発見メッセージを処理し、応答メッセージを遅い時間スタンプを有する発見メッセージを発したノード、例えばノード130へ返信する。そして、最小コストマルチキャストパスがノード105からノード120へ確立された場合、ノード130は、ランダムな(または所定の)時間待機した後に発見メッセージを発し、上述したように先に進む。

【0031】本発明の一実施形態において、本発明の原理

は、各マルチキャストグループに対する各ネットワークノードにおいて具現化されるステートマシンにおいて具体化される。そのようなステートマシンの一例が図3に示されており、5個の主な状態、即ちアイドル状態301、待機状態302、仮の状態303、アクティブ状態304および再試行状態305からなる。

【0032】図3について説明する前に、マルチキャストに参加を試みるノードにより送られ得る異なるメッセージおよびマルチキャストセッションに参加するための要求に応じてネットワーク中の他のノードにより送信されるメッセージについて説明する。マルチキャストツリー/セッションに参加しようとするノードは、上述した方法で、要求メッセージを送信する。

【0033】要求メッセージのヘッダは、とりわけ、マルチキャスト情報のアイデンティティ、送信ノードのIDおよび時間スタンプを含む。また、これは、再試行のカウントを含むことができる。要求メッセージは、アイドル状態にあるノードによってのみ開始され得る。1つのノードが、要求メッセージの受信に応じて、応答メッセージを送信する。

【0034】応答メッセージは、要求メッセージのヘッダ情報および、現在のノードからツリー上にすでにある最も近いノードへの最短パスのコストを識別する関連づけられたコストパラメータを含む。このコストは、応答が1つのノードから別のノードへ移動する時に、更新される。このコストは、特定のパス上にアクティブノードがない場合に、-1にセットされる。

【0035】1つのノードは、上述したように、別のリクエストが既に処理されたと決定された場合に、再試行メッセージを送信する。早い時間スタンプを伴う要求は、上述したように、再試行メッセージが遅い時間スタンプを有する要求の開始ノードへ送信される間、継続することを許容される。1つのノードは、マルチキャストセッション/ツリーに参加することを望む場合に、参加要求メッセージを送信する。

【0036】最も近いノードへのパスは、このメッセージ内で符号化される。アクティブノードは、参加要求メッセージの受信に応じて、参加肯定応答メッセージを送信する。このメッセージの受信は、1つのノードがマルチキャストツリー上にあることを示す。ノードは、マルチキャストセッション/ツリーを離れることを望む場合に、アクティブな隣のノードへ分離メッセージを送信する。

【0037】図3において、アイドル状態において、ノードは、マルチキャスト情報（即ちツリー）に属する情報を有しない。ノードは、要求メッセージを送信した後に、待機状態に入り、応答を待つ。ノードは、参加要求メッセージをマルチキャストツリー上に既にある最も近いノードに向かって送信した後に、仮の状態に入る。この状態において、ノードは、参加肯定応答メッセージを

待つ。

【0038】ノードは、マルチキャストセッションに参加した場合、即ちマルチキャストツリー上にある場合、アクティブ状態に入る。ノードは、既にツリー上にあるノードから参加肯定応答メッセージを受信した場合に、この状態に到達する。また、アクティブ状態にあるノードは、それに付随しかつツリーに属する全てのリンクに関するツリー情報を維持することができる。しかし、そのノードは、ツリーに参加するために、新しい要求に関するいかなる状態情報も維持しない。

【0039】ノードは、マルチキャストツリーに参加するために、リクエストの開始ノードである場合に、再試行状態に達し、ツリーのリンクの内の1つを経由して、再試行メッセージを受信する。この状態において、ノードは、他の要求に関するいかなる状態情報も維持することを必要としない。また、このノードは、元の要求を再送信する前に、ランダムな時間を待機する。このノードは、送信する要求メッセージの数を追跡することもできる。

20 【0040】アイドル状態の拡張したバージョンが、図4に示されている。上述したように、ノードは、アイドル状態にある場合に、要求メッセージを発する。

【0041】この状態にあるノードのみが、要求メッセージを発することができる。要求メッセージは、タプル<送信ノードID、時間スタンプ、再試行カウント>を使用して識別される。ここで、送信ノードIDは、メッセージの発信者であるノードのIDであり、時間スタンプは、送信IDにおけるカウンタ値であり、再試行カウントは、ノードがツリーに参加した試みの数である。再試行カウントは、初期には0である。

30 【0042】初期の要求メッセージは、開始ノードに根を有する最短パスツリーを経由して全てのリンク上に送られる。そして、ノードは、待機状態に入る。ノードが、要求メッセージ（アクションパス401）または、参加要求メッセージ（アクションパス402）のいずれかを受信したときに、以下の動作が行われる。例えば、要求メッセージがノードBから受信され、したがってタプル<ID_B, CNT_{R_B}, RC_B>を含むと仮定する。そして、受信ノードは以下の動作を行う。

40 【0043】メッセージが最短パス上にない場合、コスト=-1で応答メッセージを開始ノードに送信する。そうでない場合、時間スタンプ（ローカルカウンタ）を更新する。メッセージの状態情報を保存する。要求メッセージを送る。状態を待機状態に変化させる。利用可能なリンクがない場合、コスト=-1とともに応答メッセージを送信する。状態をアイドル状態に変化させる。

【0044】参加要求メッセージがノードBから受信された場合、受信ノードは、以下のように処理を進める。

50 【0045】ジョインリクエストメッセージを次のノ

11

ードへ送る。状態を仮の状態へ変化させる。現在のノードが、メッセージ中に記録された最終的な宛先である場合、応答メッセージを開始ノードへ返信する。この状態は、ノードが、要求メッセージの送信とノードBへの参加要求メッセージの送信との間の時間に、その状態をアクティブ状態からアイドル状態へ変化させる場合に、起こり得る。

【0046】待機状態の拡張バージョンが、図5A、5Bおよび6に示されている。図5Aおよび図5Bに対して、ノードが、タプル<ID_A, CNTR_A, RC_A> 10を含む要求メッセージをノードAから受信したと仮定する。

【0047】またノードが、タプル<ID_B, CNTR_B, RC_B>を含む要求メッセージをノードBから受信した場合、ノードは、以下のアクションをとる(パス501)。

【0048】ローカル時間スタンプカウンタを更新する。メッセージが最短パスを經由して受信されなかった場合、応答メッセージをノードBへ送信し、コストを1にセットする。メッセージが、最短パスを經由して受信され、ノードBがノードAよりも時間的に優先する場合、再試行メッセージをノードAに送信する。ノードAに属する状態情報をクリアする。ノードBに属する情報を保存し、要求メッセージを送る。

【0049】ノードがそのリンクの内の1つを經由して応答メッセージを受信した場合、そのノードは、以下のアクションをとる(パス502)。

【0050】応答メッセージが状態情報を含んでいない場合、これを無視する。最後の応答でない場合、コスト情報を格納し、他の応答メッセージを待つ。最後の応答メッセージである場合、(メッセージのコストフィールド+リンクコスト)を最小化することにより最適なメッセージを決定する。そして、コストを更新した後に、要求の開始ノードに向かって最適メッセージを送る。状態をアイドル状態に変化させる。要求メッセージの開始ノードである場合、最短ノードへの最短パスを計算する。参加要求メッセージを最短パスを經由してそのノードへ送信する。状態を仮の状態に変化させる。

【0051】そのノードがそのリンクの内の1つを經由して再試行メッセージを受信した場合、そのノードは、 40以下のアクションをとる(図6のパス601)。

【0052】再試行メッセージが状態情報を含んでいない場合、それを無視する。状態情報をクリアし、再試行メッセージを要求の開始ノードへ送る。状態をアイドル状態に変化させる。要求の開始ノードである場合、状態を再試行状態に変化させる。再試行カウント値をインクリメントする。ランダムな時間待機し、更新された再試行カウント値と共に要求メッセージを再送信する。ここで、同じカウント値(CNTR)が、公正を保証するために、使用されることになる。

12

【0053】このノードが、参加要求メッセージをあるノード、例えばノードBから受信した場合、そのノードは以下のように処理を進める(図6のパス602)。

【0054】状態を仮の状態に変化させる。次のノード情報が利用可能である場合、参加要求メッセージを次のノードへ送り、再試行メッセージをノードAに送信する。次のノードがない場合、再試行メッセージを参加要求メッセージの開始ノードへ送信する。

【0055】図3の仮の状態の拡張バージョンが、図7に示されており、ノードが、別のノード、例えばノードAにより開始された参加要求メッセージを受信した後、仮の状態に入る。その時点において、ノードがタプル<ID_B, CNTR_B, RC_B>を含むノードBから要求メッセージを受信した場合、受信ノードは、以下の方法で処理を進める(パス701)。

【0056】要求メッセージが、最短パスを經由して受信されなかった場合、コスト=-1とともに応答メッセージを開始ノードへ返信する。そうでない場合、ローカルカウンタ値を更新する。再試行メッセージをメッセージの開始ノードへ送信する。

【0057】一方、そのノードが再試行メッセージを受信した場合、そのノードは以下のアクションをとる(パス702)。

【0058】再試行メッセージが参加要求メッセージと一致した場合、再試行メッセージをノードAへ送り返す。状態をアイドル状態に変化させる。再試行メッセージが参加要求メッセージと一致しない場合、これを無視する。

【0059】そのノードが、別の方法で参加肯定応答メッセージを受信した場合、そのノードは、以下のアクションをとる(パス703)。

【0060】状態をアクティブ状態に変化させる。参加肯定応答メッセージをノードAに送る。全ての状態情報をクリアする。参加要求メッセージが受信されたリンクを含むように、マルチキャストツリー情報を更新する。

【0061】図3のアクティブ状態の拡張されたバージョンが、図8に示されており、ノードは、マルチキャストツリー上にすでにある。この場合において、そのノードのみが、新しい要求および参加要求に応答する。他のメッセージは、無視され得る。

【0062】アクティブ状態にあるノードが、タプル<ID_B, CNTR_B, RC_B>を含むノードBから要求メッセージを受信した場合、受信ノードは、以下のアクションをとる(パス801)。

【0063】要求メッセージが、最短パスを經由して受信されなかった場合、コスト=-1と共に、応答メッセージを開始ノードへ送信する。そうでない場合、ローカルカウンタ値を更新する。コスト=0と共に応答メッセージを送信する。

50 【0064】一方、そのノードが、参加要求メッセージ

をノードBから受信した場合、そのノードは、参加肯定応答メッセージをノードBへ返信する(パス802)。そのノードが、その他の方法で分離メッセージをノードBから受信した場合、そのノードは、以下のアクションをとる。

【0065】マルチキャストツリー情報を更新する。アクティブな隣のノードが1つのみであり、かつそのノードがマルチキャストグループのメンバーでない場合、分離メッセージを隣のノードへ送信する。状態をアイドル状態に変化させる。

【0066】再試行状態の拡張されたバージョンが、図9に示されている。ノードは、要求メッセージを送信した後に、再試行状態に入る。ノードAが再試行状態にあり、関連づけられたタプル情報が<ID_A, CNTR_A, RC_A>であり、ノードAが要求メッセージをそのタプル情報が<ID_B, CNTR_B, RC_B>であるノードBから受信すると仮定する。この場合に対して、ノードAは、以下のアクションをとる(パス901)。

【0067】メッセージが最短パス上にない場合、-1にセットされたコストとともに応答メッセージを返信する。ローカルカウンタを更新する。ノードAが時間的にノードBに優先する場合、再試行メッセージをノードBに送信する。ノードBが時間的にノードAに優先する場合、ノードBの状態情報を保存し、要求メッセージを送る。ノードAの情報を保存し、ノードAの再試行タイマーを停止する。ノードAは、ノードBが参加するまで、マルチキャストグループに参加することができない。状態を待機状態に変化させる。(ノードAはノードBを待つことになる)。

【0068】一方、参加要求メッセージが、ノードBから受信された場合、そのノードは以下のアクションをとる(パス902)。

【0069】状態を仮の状態に変化させる。参加要求メッセージをパス上の次のノードへ送る。ノードAの情報を保存し、待機する。(ノードAはノードBがマルチキャストグループに参加するまで、要求メッセージを再送信することができない)。再試行タイマーが時間満了となると、再試行カウント値をインクリメントし、タプル<ID_A, CNTR_A, RC_A>を含む要求メッセージを送信する。CNTR_Aの古い値は、公正を保証するために使用される。

【0070】以上は、複数のいわゆるピアグループから形成されるネットワークにおいて、容易に具現化され得る。ピアグループは、より小さなネットワークを形成するネットワークノードのグループである。特に、多数のノードから形成される大規模ネットワークは、数の効率を向上させるために論理的に再構成され得る。そのような再構成は、複数のノードをそれぞれがより小さなネットワークとして動作するグループにすることを必要とする。ここで、接続管理機能は、ネットワーク階層の各レ

ベルに論理的に拡張される。

【0071】状態情報は、各物理ノードにおいて維持される。また、各グループは、ピアグループリーダー(PGL)によりネットワーク階層のより高いレベルにおける論理ノードとして表現される。PGLは、その階層レベルにおける論理ノードに対する別個の状態情報を維持する。論理ノードは、ネットワーク階層のより低いレベルにおける物理ノードと同じ状態遷移を実行する。この意味において、我々の新規なプロトコルは、以下のような修正(拡張)を伴うそのようなネットワークに適用可能である。

【0072】発見段階

i) 参加ノードは、そのPGLが階層の次に高いレベルにおける「発見」段階に入ることを要求する。論理レベルにおいて、論理ノード間の予め確立された仮想接続が、要求メッセージおよび応答メッセージを送信するために使用される。発見段階において、論理ノードは、物理ノードと同様に、アイドル状態から待機状態および再試行状態への遷移をなし得る。

20 【0073】ii) PGLが高レベルにおいて既にアクティブである(これは、いくつかの他のアクティブなノードがピアグループ中に存在することを意味し得る)場合、この情報は、要求ノードへ返信される。この情報の受信により、要求ノードは、ピアグループ内の最も近いアクティブノードを決定するためのプロトコルを実行する。物理的または論理的な最も近いアクティブノードのIDが、低レベルの要求ノードへ返信される。現在のレベルが、最も低いレベルである場合、要求ノードは、参加段階に入る。

30 【0074】iii) そうでなければ、PGLは、それが階層の最上レベルに到達するまでより高いレベルにおいて上述のステップを反復して実行する。そのレベルにおいて、PGLは、発見段階に入る。発見段階の終わりにおいて、要求ノードは、最も近いアクティブノードのIDを有する。このIDは、同じピアグループ(単一のピアグループの場合)内のノードの物理的IDまたは複数のピアグループの場合、論理的ピアグループ中のより高いレベルのノードの論理的IDのいずれかであり得る。

40 【0075】参加段階

利用可能な地勢的情報に基づいて、要求ノードは、最も近いノードへのパスを決定する。このパスは、最も近いノードが、より高いレベルの論理ピアグループ中の論理ノードである場合、完成されない。決定されたパスは、指定遷移リスト(DTL)の形式で格納される。これは、ATMフォーラム、2570 West, El Camino Real, Suite 304, Mountain View, California 94040-1313, から入手可能な私設網ノードインタフェース(PNNI)仕様バージョン1.0に説明されている。DTLは、最も近いノードへ送られる前に、参加要求メッセージ中に

埋め込まれる。参加要求メッセージの受信により、ノードは以下のプロトコルを実行する。

【0076】i) そのノードが既にアクティブである場合、参加肯定応答メッセージを返信する。参加肯定応答メッセージが、逆方向において同じパスで開始ノードへ返信される。

【0077】i i) そうでない場合、DTLスタックが空でない場合、参加要求メッセージは、DTLにおいて利用可能な情報に基づいて、次のノードへ送られる。新しいパスは、上述した参考文献に説明されているように、10 入り口ノードによりDTLへ追加され得る。また、参加要求メッセージおよび参加肯定応答メッセージは、ネットワーク信号チャネル上を送信され得る。参加要求メッセージまたは参加肯定応答メッセージは、境界リンクを通して送信され、適切なレベルにおける論理ノードは、その論理ノードがそれぞれ仮の状態またはアクティブ状態への状態遷移をなし得るように、通知されなければならない。これは、参加要求または参加肯定応答メッセージのコピーを適切な論理ノードへ送信することにより行われる。

【0078】i i i) DTLスタックが空の場合、そのノードは、最も近いアクティブノードを決定するために、発見段階に入る。そして、最も近いノードの情報は、DTLとして符号化され、そのメッセージは最も近いノードへ送信される。

【0079】参加しているノードがマルチキャストグループを離れる時に、同様なアクションがとられる。この場合において、分離メッセージは、参加ノードが唯一のアクティブな隣のノードを有する場合に、アクティブな隣のノードへ送信される。そして、参加ノードは、アイドル状態へ入る。隣のノードは、参加ノードではなく、唯一のアクティブな隣のノードを有する場合、分離メッセージをその最も近い隣のノードへ送る。分離メッセージが、境界リンクを通して、送信される場合、メッセージのコピーは、ノードの全体的な状態がアイドル状態へ変化され得るように、適切な論理ノードへ送信される。

【0080】直前に説明された拡張プロトコルは、図10乃至13に示された例との関係で説明することにより、さらに理解されるであろう。ネットワークノード（例えばATM交換機）の3個のピアグループ201、202および203から形成されたネットワークを示す。ピアグループ(a)201中のノードは、それぞれA. 1乃至A. 5の符号がつけられており、ピアグループ(b)202中のノードは、それぞれB. 1乃至B. 5の符号がつけられており、ピアグループ(c)203中のノードは、それぞれC. 1乃至C. 5の符号がつけられている。

【0081】黒い(塗りつぶされた)ノードは、図10乃至13のそれぞれにおいて、各ピアグループに対するPGLとして働く。図10に対して、ノードA. 3は、

マルチキャストグループに参加することを望み、論理PGLA、BおよびCを含む楕円で示されているより高いレベルにおける発見段階にPGLを入らせる、要求をそのPGL(ノードA)へ送信する。PGLAが最も高いレベルに入るので、どれが要求されたマルチキャスト情報を有するかを決定するために、そのピアグループの他のメンバーへ、要求メッセージを送信する。

【0082】PGLAが、ピアグループBおよびC中にマルチキャストのメンバーがないことを決定したと仮定する。この情報は、ノードA. 3へ送られ、ノードA. 3の状態はアクティブになる。マルチキャストツリーは、この時点で、単一のノードA. 3から構成される。全体的に、論理ノードAのステートマシンは、アクティブ状態への遷移を生じる。各レベルにおけるアクティブ状態は、図11中のボックスで示されている。各レベルにおける参加ノードは、影付きボックスで示されている。

【0083】ノードC. 4は、マルチキャストグループへの参加を望み、発見段階に入るために、そのPGL(ノードC)へ要求を送ると仮定する。これに20 応答して、ノードCは、要求メッセージを論理ノードBおよび論理ノードAへ送信する。ノードCは、論理ノードBからの同時の参加要求がある場合、再試行状態に入る可能性がある。他の要求がない場合、論理ノードAのみがアクティブであり、この情報は、論理ノードCにより、ノードC. 4へ送信される。このアクションは、ノードC. 4に対する発見段階を完了させる。

【0084】参加段階において、ノードC. 4は、ピアグループAへのパスをDTLとして符号化し、参加要求メッセージをピアグループAへ送る。このメッセージは、境界リンク(C. 2-B. 4)を通してピアグループBに入る。ここで、ノードB. 4は、入り口ノードであり、DTLスタックは、空ではない(参加段階のステップ(i i))。ノードB. 4は、そのピアグループを通るピアグループAへのパスを識別し、そのパスへDTLを付加する。そして、そのメッセージをピアグループAへ送る。

【0085】ノードB. 4は、参加要求メッセージのコピーもPGLノードB. 1へ送信する。このPGLノードB. 1は、論理ノードBのための全体的ステートマシンを維持する。ノードB. 1は、ノードBの全体的ステートマシンを仮の状態へ更新する。このメッセージがリンク(B. 5-A. 5)を通してピアグループAに入ると仮定する。DTLスタックは、今、空である(参加段階のステップ(i i i))。

【0086】ノードA. 5は、メッセージの受信により、メッセージを受信すべき最も近いノード、即ちノードA. 3を決定するために、発見段階に入る。そして、ノードA. 3へのパスをDTLに付加し、参加要求メッセージをノードA. 3へ送信する。アクティブ状態にあ

るノードA. 3は、参加肯定応答メッセージを返信することにより、メッセージの受信に対して応答する（参加段階のステップ（i））。参加肯定応答メッセージは、参加要求メッセージがたどったパスを再度たどる。

【0087】参加肯定応答メッセージが、境界リング（A. 5—B. 5）を通るとき、論理ノードBのステートマシンは、アクティブ状態に更新される。論理ノードCは、参加肯定応答メッセージがリンク（B. 4—C. 2）を通して送信されるときに、アクティブになる。結果として得られるマルチキャストツリーは、図2

に点線で示されている。ノードB. 2は、図13に示されているように、ノードB. 4を参加させることによりマルチキャストツリーに参加することができる。

【0088】参加ノードA. 3が、マルチキャストグループを離れる場合、分離メッセージをノードB. 5へ伝播するノードA. 5へ分離メッセージを送信する。また、ノードA. 5は、論理ピアグループAの全体的ステートマシンがアイドル状態へ変化され得るように、分離メッセージのコピーをノードA. 1へ送信する。また、ノードB. 5は、分離メッセージをノードB. 4へ送信する。しかし、ノードB. 4は2つのアクティブな隣のノード、即ちノードB. 1およびC. 2を有するので、後者のメッセージの伝播は、ノードB. 4において停止する。得られるツリーは、図14に示されている。

【0089】今、ノードB. 2が同様にマルチキャストグループを離れると仮定する。この時点において、分離メッセージはノードC. 4へ伝播する。分離メッセージがリンクB. 4—C. 2を通して送信される場合、ピアグループに対する全体的ステートマシンがアイドル状態へ変化され得るように、コピーもノードB. 1へ送信される。ノードC. 4のみからなる最終的なツリーが図15に示されている。

【0090】

【発明の効果】以上説明したように、本発明によれば、安価かつ高効率のマルチキャストグループを探し出し、これに参加し、かつこれを離れるためのマルチキャストプロトコルを提供することができる。

【図面の簡単な説明】

【図1】本発明の原理が実行される通信ネットワークの構成を示す図。

【図2】図1のノードのうちの特定の一つに根を有するルーティングツリーを示す図。

【図3】本発明の原理によるノードの動作を示す状態図。

【図4】図3に示された動作状態のそれぞれの拡張バージョンを示す図。

【図5】図3に示された動作状態のそれぞれの拡張バージョンを示す図。

【図6】図3に示された動作状態のそれぞれの拡張バージョンを示す図。

【図7】図3に示された動作状態のそれぞれの拡張バージョンを示す図。

【図8】図3に示された動作状態のそれぞれの拡張バージョンを示す図。

【図9】図3に示された動作状態のそれぞれの拡張バージョンを示す図。

【図10】例示的な階層ネットワークにおける本発明の原理の動作を示す図。

【図11】例示的な階層ネットワークにおける本発明の原理の動作を示す図。

【図12】例示的な階層ネットワークにおける本発明の原理の動作を示す図。

【図13】例示的な階層ネットワークにおける本発明の原理の動作を示す図。

【図14】例示的な階層ネットワークにおける本発明の原理の動作を示す図。

【図15】例示的な階層ネットワークにおける本発明の原理の動作を示す図。

【符号の説明】

100 ネットワーク

105, 110, 115, 120, 125, 130, 1

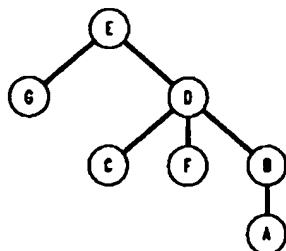
35 ノード

150 マルチメディア事象

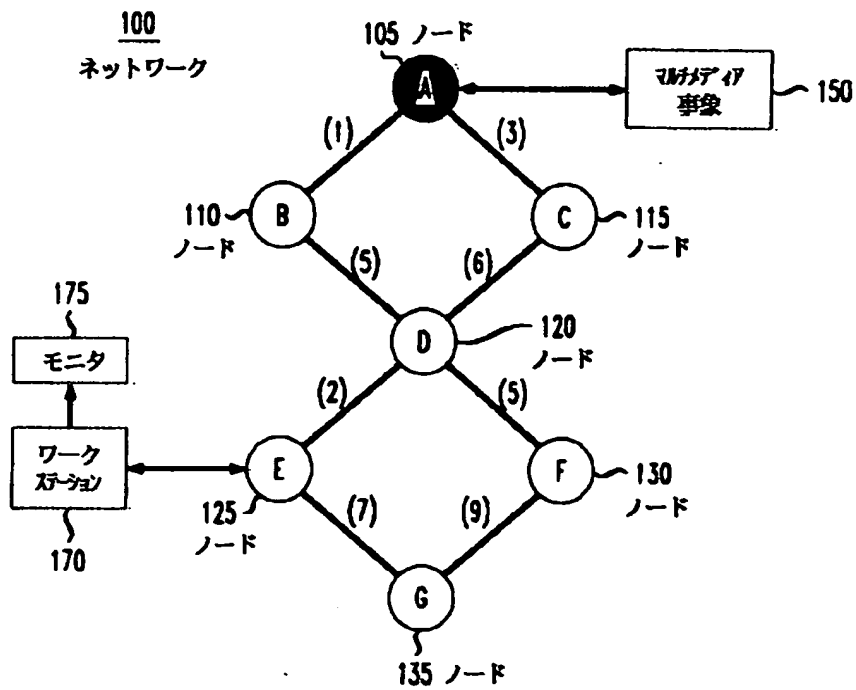
170 ワークステーション

175 モニタ

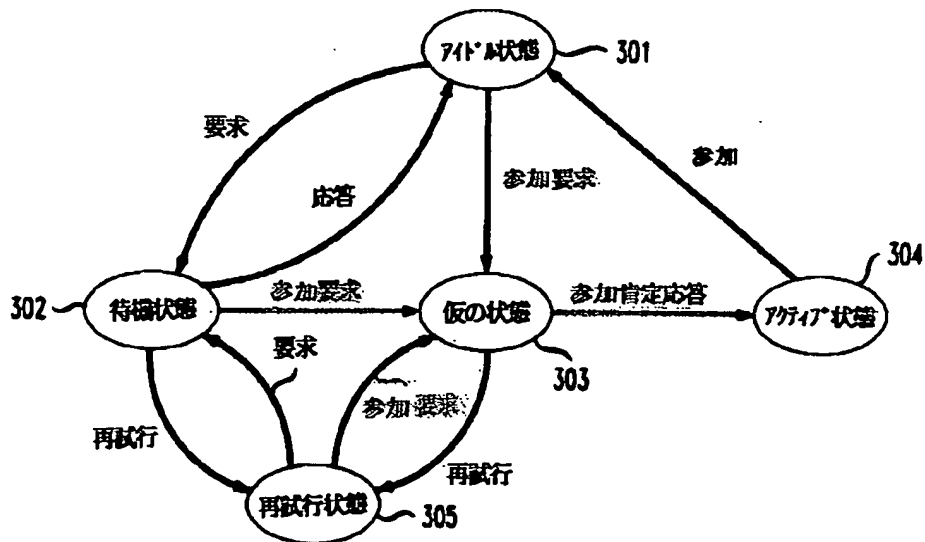
【図2】



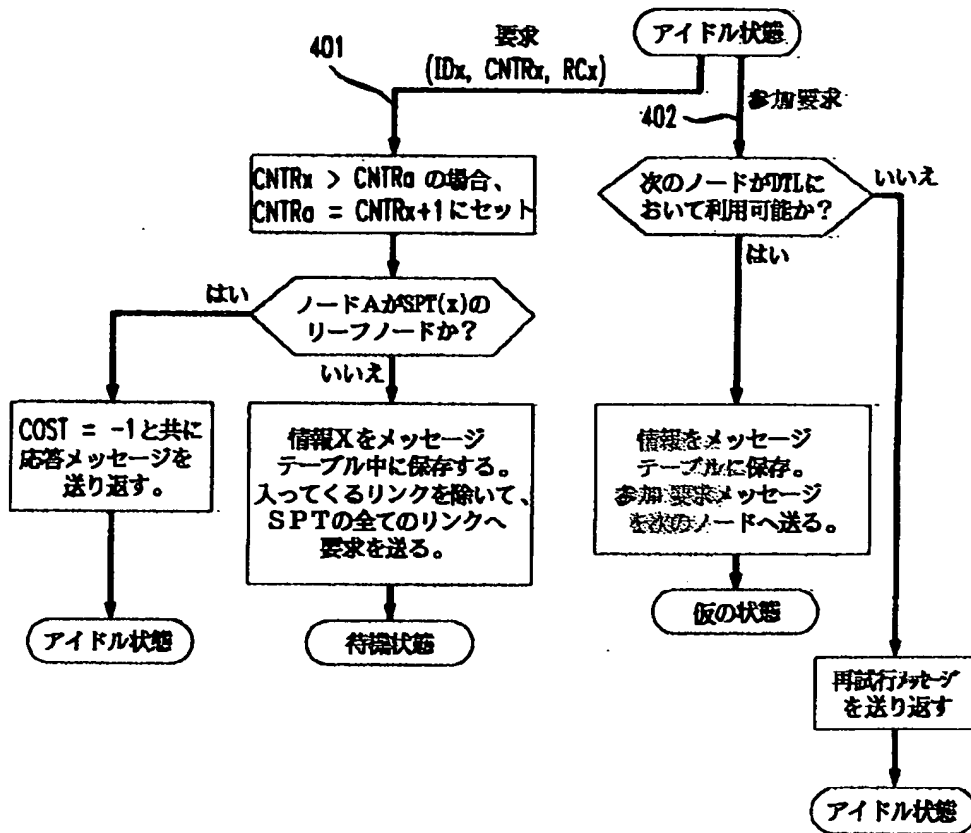
【図1】



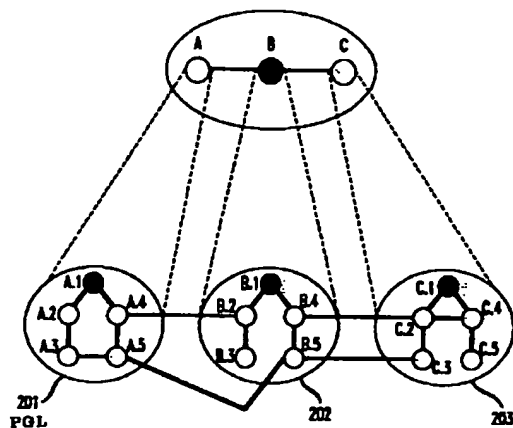
【図3】



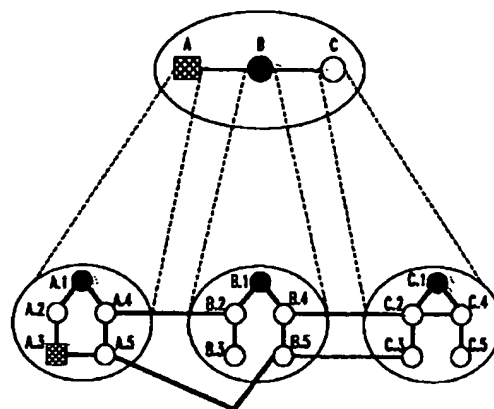
【図4】



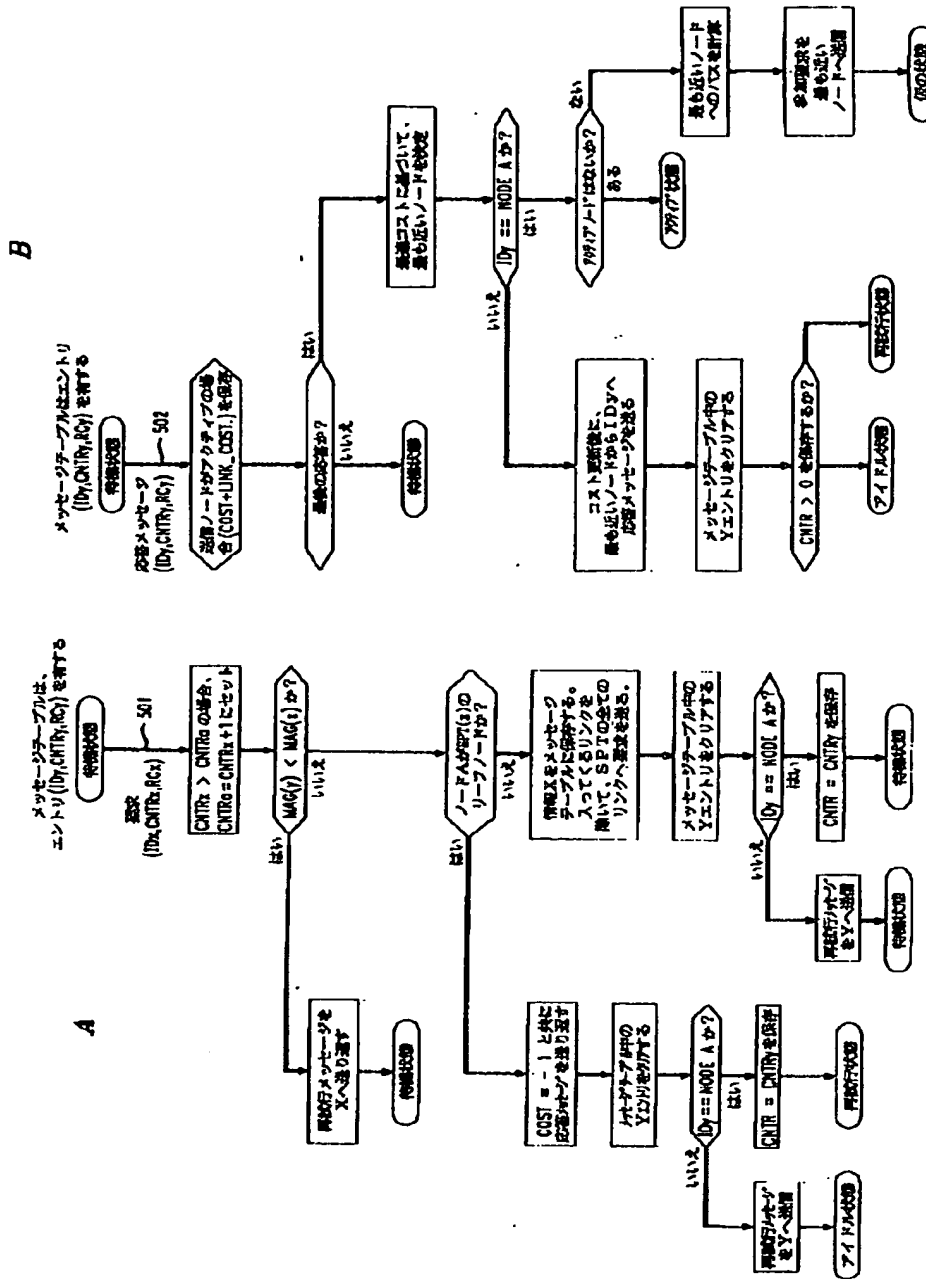
【図10】



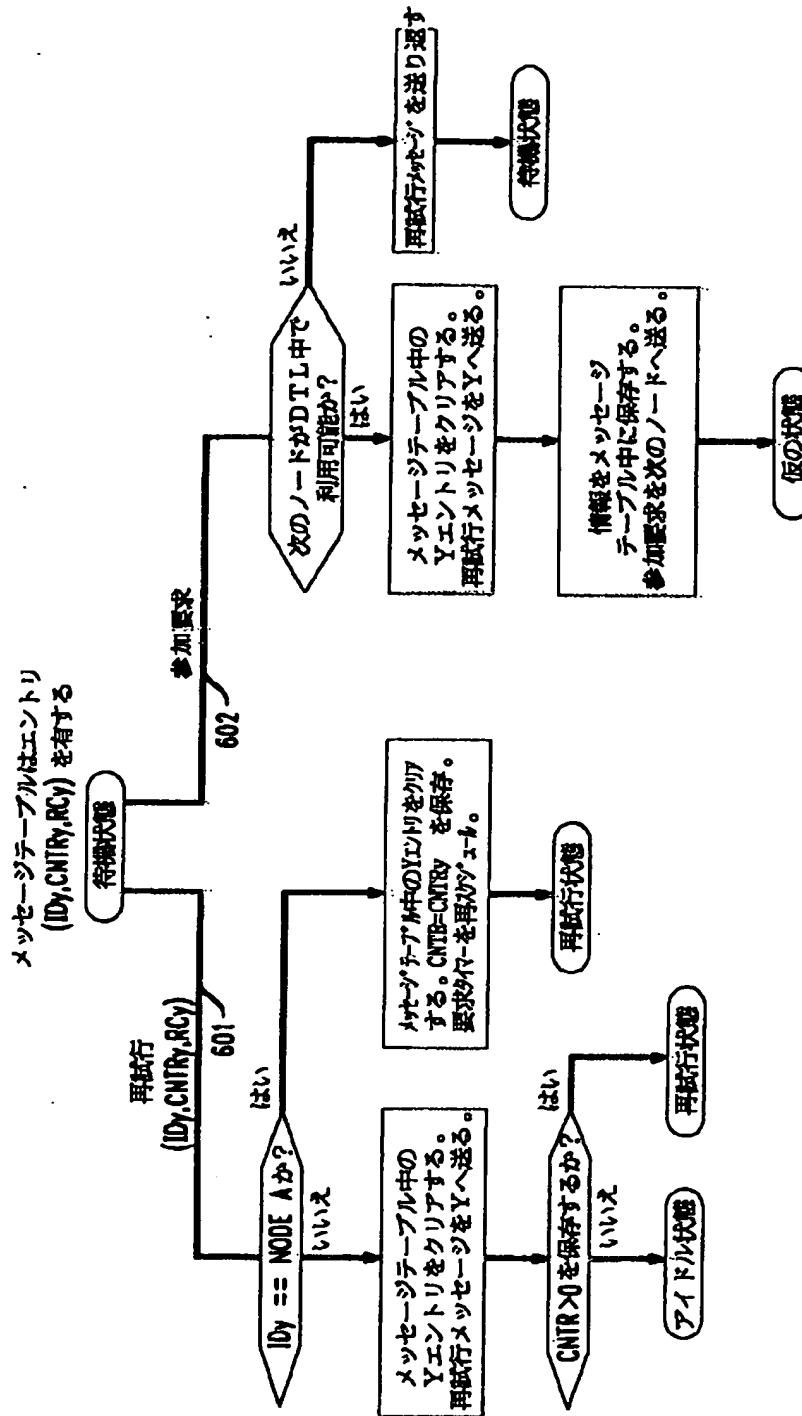
【図11】



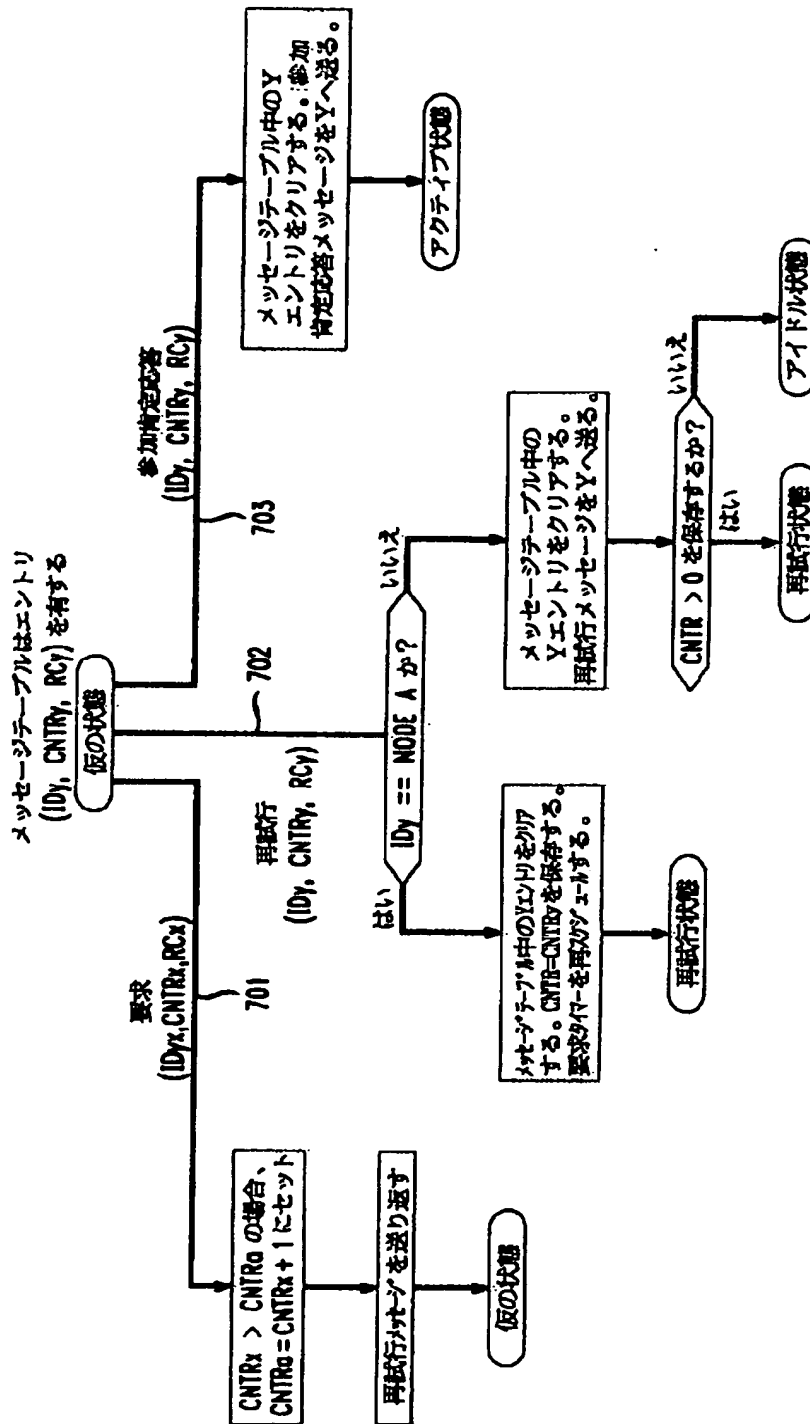
【图5】



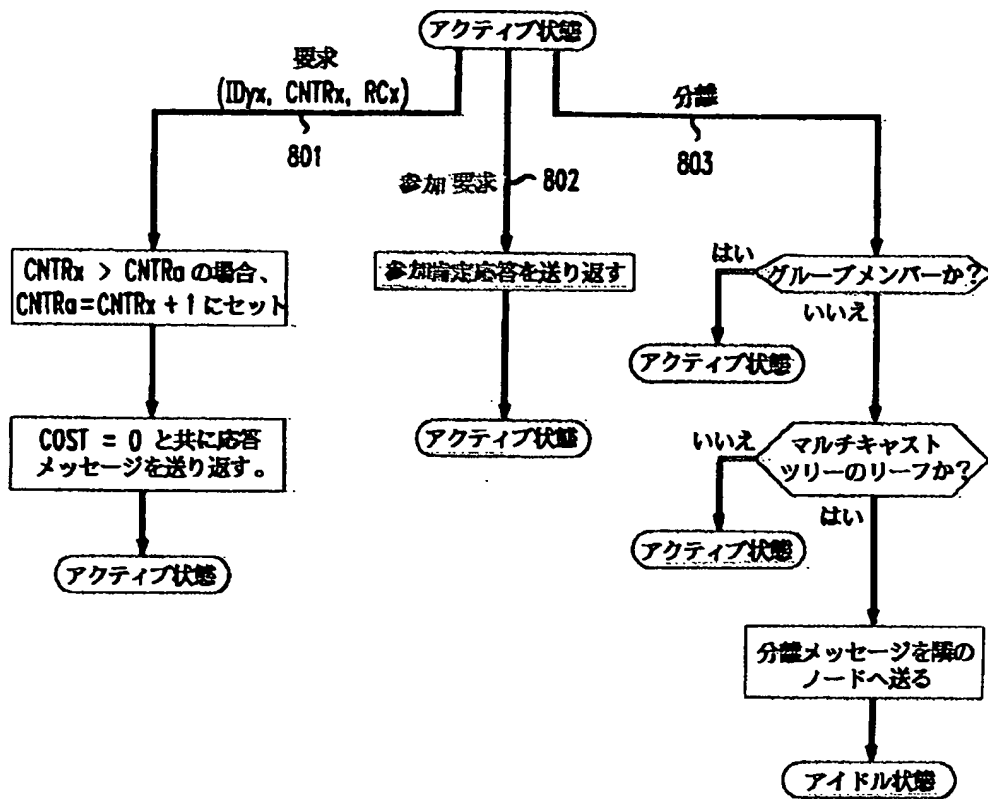
【図6】



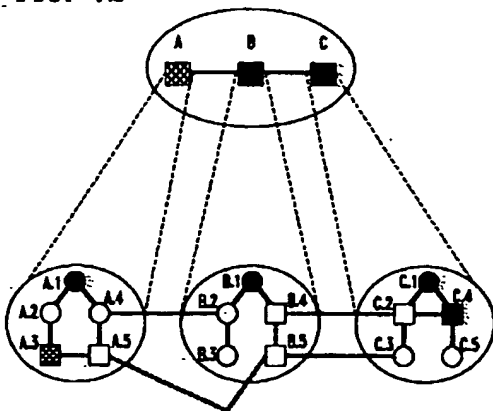
【図7】



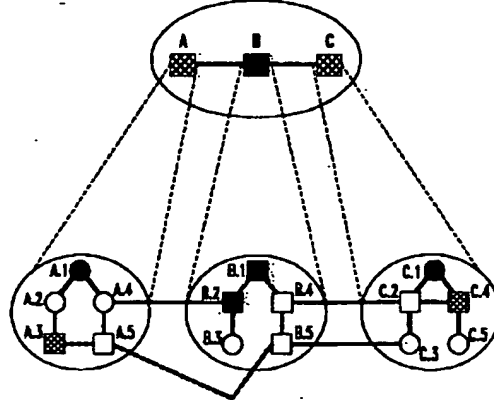
【図8】



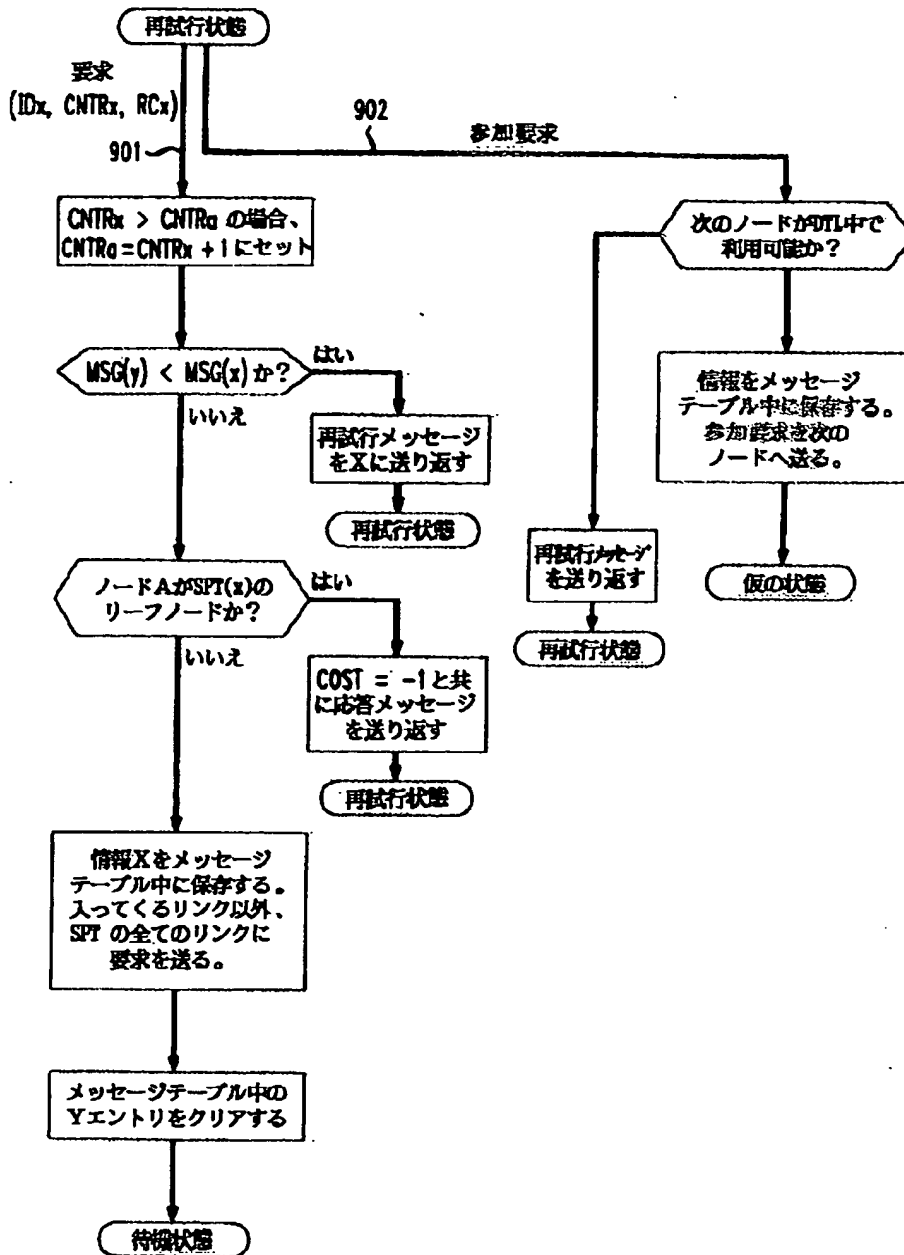
【図12】



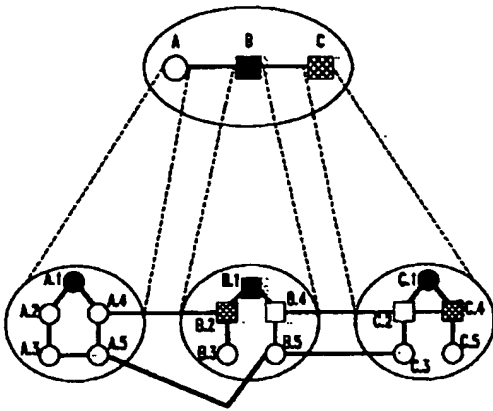
【図13】



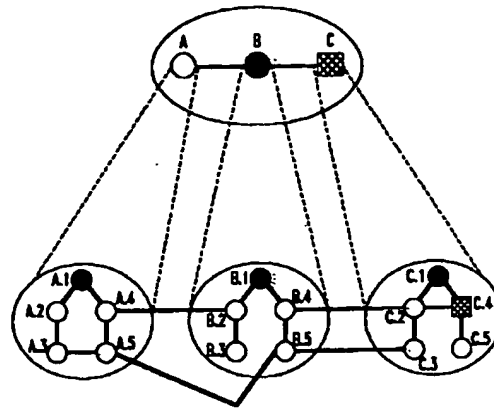
【図9】



【図14】



【図15】



フロントページの続き

(71)出願人 596077259
600 Mountain Avenue,
Murray Hill, New Je
rsey 07974-0636 U. S. A.

(72)発明者 ビジェイ ボチャンバリ クマー
アメリカ合衆国、07728 ニュージャージ
ー、フリーホールド、タワー ロード 3

(72)発明者 カウリギ スリニバサ ラガベンドラ
アメリカ合衆国、99163 ワシントン、ア
ルマン、エッジ ノール ドライブ 829
エス. イー.

(72)発明者 ラマナサン ベンカデスワラン
アメリカ合衆国、07733 ニュージャージ
ー、ホルムデル、サウス ホルムデル ロ
ード 14ビー

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.